

0409

A2



OIPE

RAW SEQUENCE LISTING

DATE: 04/11/2002

PATENT APPLICATION: US/10/020,139

TIME: 09:17:39

Input Set : N:\Crf3\RULE60\10020139.raw

Output Set: N:\CRF3\04112002\J020139.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
5 (i) APPLICANT: DUAN, ROXANNE
6 RUBEN, STEVEN
8 (ii) TITLE OF INVENTION: Parotid Secretory Protein
10 (iii) NUMBER OF SEQUENCES: 18
12 (iv) CORRESPONDENCE ADDRESS:
13 (A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
14 (B) STREET: 9410 KEY WEST AVENUE
15 (C) CITY: ROCKVILLE
16 (D) STATE: MD
17 (E) COUNTRY: US
18 (F) ZIP: 20850
20 (v) COMPUTER READABLE FORM:
21 (A) MEDIUM TYPE: Floppy disk
22 (B) COMPUTER: IBM PC compatible
23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
26 (vi) CURRENT APPLICATION DATA:
C--> 27 (A) APPLICATION NUMBER: US/10/020,139
C--> 28 (B) FILING DATE: 18-Dec-2001
29 (C) CLASSIFICATION:
31 (vii) PRIOR APPLICATION DATA:
32 (A) APPLICATION NUMBER: US/08/993,529
33 (B) FILING DATE:
36 (viii) ATTORNEY/AGENT INFORMATION:
37 (A) NAME: BROOKES, ANDERS A.
38 (B) REGISTRATION NUMBER: 36,373
39 (C) REFERENCE/DOCKET NUMBER: PF348
41 (ix) TELECOMMUNICATION INFORMATION:
42 (A) TELEPHONE: (301) 309-8504
43 (B) TELEFAX: (301) 301-8439
46 (2) INFORMATION FOR SEQ ID NO: 1:
48 (i) SEQUENCE CHARACTERISTICS:
49 (A) LENGTH: 1028 base pairs
50 (B) TYPE: nucleic acid
51 (C) STRANDEDNESS: single
52 (D) TOPOLOGY: linear
54 (ii) MOLECULE TYPE: DNA (genomic)
57 (ix) FEATURE:
58 (A) NAME/KEY: CDS
59 (B) LOCATION: 49..795
61 (ix) FEATURE:

ENTERED

RAW SEQUENCE LISTING

DATE: 04/11/2002

PATENT APPLICATION: US/10/020,139

TIME: 09:17:39

Input Set : N:\Crif3\RULE60\10020139.raw

Output Set: N:\CRF3\04112002\J020139.raw

```

62      (A) NAME/KEY: sig_peptide
63      (B) LOCATION: 49..100
65      (ix) FEATURE:
66      (A) NAME/KEY: mat_peptide
67      (B) LOCATION: 103..795
70      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
72 CACGAGATTT CATGAGCATC CTCCTCTAAA CGCGTGTCAG GACAAAAG ATG CTT CAG      57
73                                     Met Leu Gln
74                                     -18
76 CTT TGG AAA CTT GTT CTC CTG TGC GGC GTG CTC ACT GGG ACC TCA GAG      105
77 Leu Trp Lys Leu Val Leu Leu Cys Gly Val Leu Thr Gly Thr Ser Glu
W--> 78 -15          -10          -5          1
80 TCT CTT CTT GAC AAT CTT GGC AAT GAC CTA AGC AAT GTC GTG GAT AAG      153
81 Ser Leu Leu Asp Asn Leu Gly Asn Asp Leu Ser Asn Val Val Asp Lys
82          5          10          15
84 CTG GAA CCT GTT CTT CAC GAG GGA CTT GAG ACA GTT GAC AAT ACT CTT      201
85 Leu Glu Pro Val Leu His Glu Gly Leu Glu Thr Val Asp Asn Thr Leu
86          20          25          30
88 AAA GGC ATC CTT GAG AAA CTG AAG GTC GAC CTA GGA GTG CTT CAG AAA      249
89 Lys Gly Ile Leu Glu Lys Leu Lys Val Asp Leu Gly Val Leu Gln Lys
90          35          40          45
92 TCC AGT GCT TGG CAA CTG GCC AAG CAG AAG GCC CAG GAA GCT GAG AAA      297
93 Ser Ser Ala Trp Gln Leu Ala Lys Gln Lys Ala Gln Glu Ala Glu Lys
94 50          55          60          65
96 TTG CTG AAC AAT GTC ATT TCT AAG CTG CTT CCA ACT AAC ACG GAC ATT      345
97 Leu Leu Asn Asn Val Ile Ser Lys Leu Leu Pro Thr Asn Thr Asp Ile
98          70          75          80
100 TTT GGG TTG AAA ATC AGC AAC TCC CTC ATC CTG GAT GTC AAA GCT GAA      393
101 Phe Gly Leu Lys Ile Ser Asn Ser Leu Ile Leu Asp Val Lys Ala Glu
102          85          90          95
104 CCG ATC GAT GAT GGC AAA GGC CTT AAC CTG AGC TTC CCT GTC ACC GCG      441
105 Pro Ile Asp Asp Gly Lys Gly Leu Asn Leu Ser Phe Pro Val Thr Ala
106          100          105          110
108 AAT GTC ACT GTG GCC GGG CCC ATC ATT GGC CAG ATT ATC AAC CTG AAA      489
109 Asn Val Thr Val Ala Gly Pro Ile Ile Gly Gln Ile Ile Asn Leu Lys
110          115          120          125
112 GCC TCC TTG GAC CTC CTG ACC GCA GTC ACA ATT GAA ACT GAT CCC CAG      537
113 Ala Ser Leu Asp Leu Leu Thr Ala Val Thr Ile Glu Thr Asp Pro Gln
114 130          135          140          145
116 ACA CAC CAG CCT GTT GCC GTC CTG GGA GAA TGC GCC AGT GAC CCA ACC      585
117 Thr His Gln Pro Val Ala Val Leu Gly Glu Cys Ala Ser Asp Pro Thr
118          150          155          160
120 AGC ATC TCA CTT TCC TTG CTG GAC AAA CAC AGC CAA ATC ATC AAC AAG      633
121 Ser Ile Ser Leu Ser Leu Leu Asp Lys His Ser Gln Ile Ile Asn Lys
122          165          170          175
124 TTC GTG AAT AGC GTG ATC AAC ACG CTG AAA AGC ACT GTA TCC TCC CTG      681
125 Phe Val Asn Ser Val Ile Asn Thr Leu Lys Ser Thr Val Ser Ser Leu
126          180          185          190
128 CTG CAG AAG GAG ATA TGT CCA CTG ATC CGC ATC TTC ATC CAC TCC CTG      729

```

RAW SEQUENCE LISTING

DATE: 04/11/2002

PATENT APPLICATION: US/10/020,139

TIME: 09:17:39

Input Set : N:\Crif3\RULE60\10020139.raw

Output Set: N:\CRF3\04112002\J020139.raw

```

129 Leu Gln Lys Glu Ile Cys Pro Leu Ile Arg Ile Phe Ile His Ser Leu
130      195                200                205
132 GAT GTG AAT GTC ATT CAG CAG GTC GTC GAT AAT CCT CAG CAC AAA ACC      777
133 Asp Val Asn Val Ile Gln Gln Val Val Asp Asn Pro Gln His Lys Thr
134 210                215                220                225
136 CAG CTG CAA ACC CTC ATT TGAAGAGGAC GAATGAGGAG GACCACTGTG      825
137 Gln Leu Gln Thr Leu Ile
138      230
140 GTGCATGCTG ATTGGTTCCC AGTGGCTTGC CCCACCCCCT TATAGCATCT CCCTCCAGGA      885
142 AGCTGCTGCC ACCACCTAAC CAGCGTGAAA GCCTGAGTCC CACCAGAAGG ACCTTCCAG      945
144 ATACCCCTTC TCCTCACAGT CAGAACAGCA GCCTCTACAC ATGTTGTCCT GCCCCTGGCA      1005
146 ATAAAGGCCC ATTTCTGCAA AAA      1028
149 (2) INFORMATION FOR SEQ ID NO: 2:
151      (i) SEQUENCE CHARACTERISTICS:
152          (A) LENGTH: 249 amino acids
153          (B) TYPE: amino acid
154          (D) TOPOLOGY: linear
156      (ii) MOLECULE TYPE: protein
158      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
160 Met Leu Gln Leu Trp Lys Leu Val Leu Leu Cys Gly Val Leu Thr Gly
161 -18      -15      -10      -5
163 Thr Ser Glu Ser Leu Leu Asp Asn Leu Gly Asn Asp Leu Ser Asn Val
164      1      5      10
166 Val Asp Lys Leu Glu Pro Val Leu His Glu Gly Leu Glu Thr Val Asp
167 15      20      25      30
169 Asn Thr Leu Lys Gly Ile Leu Glu Lys Leu Lys Val Asp Leu Gly Val
170      35      40      45
172 Leu Gln Lys Ser Ser Ala Trp Gln Leu Ala Lys Gln Lys Ala Gln Glu
173      50      55      60
175 Ala Glu Lys Leu Leu Asn Asn Val Ile Ser Lys Leu Leu Pro Thr Asn
176      65      70      75
178 Thr Asp Ile Phe Gly Leu Lys Ile Ser Asn Ser Leu Ile Leu Asp Val
179      80      85      90
181 Lys Ala Glu Pro Ile Asp Asp Gly Lys Gly Leu Asn Leu Ser Phe Pro
182 95      100      105      110
184 Val Thr Ala Asn Val Thr Val Ala Gly Pro Ile Ile Gly Gln Ile Ile
185      115      120      125
187 Asn Leu Lys Ala Ser Leu Asp Leu Leu Thr Ala Val Thr Ile Glu Thr
188      130      135      140
190 Asp Pro Gln Thr His Gln Pro Val Ala Val Leu Gly Glu Cys Ala Ser
191      145      150      155
193 Asp Pro Thr Ser Ile Ser Leu Ser Leu Leu Asp Lys His Ser Gln Ile
194      160      165      170
196 Ile Asn Lys Phe Val Asn Ser Val Ile Asn Thr Leu Lys Ser Thr Val
197 175      180      185      190
199 Ser Ser Leu Leu Gln Lys Glu Ile Cys Pro Leu Ile Arg Ile Phe Ile
200      195      200      205
202 His Ser Leu Asp Val Asn Val Ile Gln Gln Val Val Asp Asn Pro Gln
203      210      215      220

```

RAW SEQUENCE LISTING

DATE: 04/11/2002

PATENT APPLICATION: US/10/020,139

TIME: 09:17:39

Input Set : N:\Crf3\RULE60\10020139.raw

Output Set: N:\CRF3\04112002\J020139.raw

205 His Lys Thr Gln Leu Gln Thr Leu Ile

206 225 230

208 (2) INFORMATION FOR SEQ ID NO: 3:

210 (i) SEQUENCE CHARACTERISTICS:

211 (A) LENGTH: 235 amino acids

212 (B) TYPE: amino acid

213 (C) STRANDEDNESS: single

214 (D) TOPOLOGY: linear

216 (ii) MOLECULE TYPE: protein

221 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

223 Met Phe Gln Leu Gly Ser Leu Val Val Leu Cys Gly Leu Leu Ile Gly

224 1 5 10 15

226 Asn Ser Glu Ser Leu Leu Gly Glu Leu Gly Ser Ala Val Asn Asn Leu

227 20 25 30

229 Lys Ile Leu Asn Pro Pro Ser Glu Ala Val Pro Gln Asn Leu Asn Leu

230 35 40 45

232 Asp Val Glu Leu Leu Gln Gln Ala Thr Ser Trp Pro Leu Ala Lys Asn

233 50 55 60

235 Ser Ile Leu Glu Thr Leu Asn Thr Ala Asp Leu Gly Asn Leu Lys Ser

236 65 70 75 80

238 Phe Thr Ser Leu Asn Gly Leu Leu Leu Lys Ile Asn Asn Leu Lys Val

239 85 90 95

241 Leu Asp Phe Gln Ala Lys Leu Ser Ser Asn Gly Asn Gly Ile Asp Leu

242 100 105 110

244 Thr Val Pro Leu Ala Gly Glu Ala Ser Leu Val Leu Pro Phe Ile Gly

245 115 120 125

247 Lys Thr Val Asp Ile Ser Val Ser Leu Asp Leu Ile Asn Ser Leu Ser

248 130 135 140

250 Ile Lys Thr Asn Ala Gln Thr Gly Leu Pro Glu Val Thr Ile Gly Lys

251 145 150 155 160

253 Cys Ser Ser Asn Thr Asp Lys Ile Ser Ile Ser Leu Leu Gly Arg Arg

254 165 170 175

256 Leu Pro Ile Ile Asn Ser Ile Leu Asp Gly Val Ser Thr Leu Leu Thr

257 180 185 190

259 Ser Thr Leu Ser Thr Val Leu Gln Asn Phe Leu Cys Pro Leu Leu Gln

260 195 200 205

262 Tyr Val Leu Ser Thr Leu Asn Pro Ser Val Leu Gln Gly Leu Leu Ser

263 210 215 220

265 Asn Leu Leu Ala Gly Gln Val Gln Leu Ala Leu

266 225 230 235

268 (2) INFORMATION FOR SEQ ID NO: 4:

270 (i) SEQUENCE CHARACTERISTICS:

271 (A) LENGTH: 235 amino acids

272 (B) TYPE: amino acid

273 (C) STRANDEDNESS: single

274 (D) TOPOLOGY: linear

276 (ii) MOLECULE TYPE: protein

281 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

283 Met Phe Gln Leu Gly Ser Leu Val Val Leu Cys Gly Leu Leu Ile Gly

RAW SEQUENCE LISTING

DATE: 04/11/2002

PATENT APPLICATION: US/10/020,139

TIME: 09:17:39

Input Set : N:\Crf3\RULE60\10020139.raw

Output Set: N:\CRF3\04112002\J020139.raw

```

284      1          5          10          15
286  Thr Ser Glu Ser Leu Leu Gly Asp Val Ala Asn Ala Val Asn Asn Leu
287      20          25          30
289  Asp Ile Leu Asn Ser Pro Ser Glu Ala Val Ala Gln Asn Leu Asn Leu
290      35          40          45
292  Asp Val Gly Ser Leu Gln Gln Ala Thr Thr Trp Pro Ser Ala Lys Asp
293      50          55          60
295  Ser Ile Leu Glu Thr Leu Asn Lys Val Glu Leu Gly Asn Ser Asn Gly
296      65          70          75          80
298  Phe Thr Pro Leu Asn Gly Leu Leu Leu Arg Val Asn Lys Phe Arg Val
299      85          90          95
301  Leu Asp Leu Gln Ala Gly Leu Ser Ser Asn Gly Lys Asp Ile Asp Leu
302      100          105          110
304  Lys Leu Pro Leu Val Phe Glu Ile Ser Phe Ser Leu Pro Val Ile Gly
305      115          120          125
307  Pro Thr Leu Asp Val Ala Val Ser Leu Asp Leu Leu Asn Ser Val Ser
308      130          135          140
310  Val Gln Thr Asn Ala Gln Thr Gly Leu Pro Gly Val Thr Leu Gly Lys
311      145          150          155          160
313  Cys Ser Gly Asn Thr Asp Lys Ile Ser Ile Ser Leu Leu Gly Arg Arg
314      165          170          175
316  Leu Pro Phe Val Asn Arg Ile Leu Asp Gly Val Ser Gly Leu Leu Thr
317      180          185          190
319  Gly Ala Val Ser Ile Leu Leu Gln Asn Ile Leu Cys Pro Val Leu Gln
320      195          200          205
322  Tyr Leu Leu Ser Thr Met Ser Gly Ser Ala Ile Gln Gly Leu Leu Ser
323      210          215          220
325  Asn Val Leu Thr Gly Gln Leu Ala Val Pro Leu
326      225          230          235

```

328 (2) INFORMATION FOR SEQ ID NO: 5:

330 (i) SEQUENCE CHARACTERISTICS:

331 (A) LENGTH: 206 amino acids

332 (B) TYPE: amino acid

333 (C) STRANDEDNESS: single

334 (D) TOPOLOGY: linear

336 (ii) MOLECULE TYPE: protein

341 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

343  Met Phe Gln Leu Gly Ser Leu Val Val Leu Cys Gly Leu Leu Ile Gly
344      1          5          10          15
346  Thr Ser Gly Ser Leu Phe Asp Ile Phe Gln Asn Pro Glu Leu Asp Val
347      20          25          30
349  Glu Ser Val Trp Ser Glu Ile Asn Tyr Arg Ile Arg Tyr Ala Leu Glu
350      35          40          45
352  Thr Met Asp Leu Asp Met Leu Ala Asp Tyr Leu Ser Lys Arg Gly Ile
353      50          55          60
355  Glu Leu Lys Ile Lys Asp Leu Arg Ile Leu Asn Leu Asn His Glu Val
356      65          70          75          80
358  Ser Pro Asn Gly Asp Glu Val Thr Leu Lys Met Pro Met Ala Leu Asn
359      85          90          95

```

VERIFICATION SUMMARY

DATE: 04/11/2002

PATENT APPLICATION: US/10/020,139

TIME: 09:17:40

Input Set : N:\Crf3\RULE60\10020139.raw

Output Set: N:\CRF3\04112002\J020139.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:78 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1